Reprinted from *Ciba Foundation Symposium on Extrasensory Perception* 1956, pp. 24–38

SOME DIFFICULTIES IN THE WAY OF SCIENTIFIC RECOGNITION OF EXTRASENSORY PERCEPTION

J. FRASER NICOL

Cambridge, Massachusetts

Reprinted from *Ciba Foundation Symposium on Extrasensory Perception* 1956, pp. 24–38

SOME DIFFICULTIES IN THE WAY OF SCIENTIFIC RECOGNITION OF EXTRASENSORY PERCEPTION

J. Fraser Nicol

Cambridge, Massachusetts

Almost three-quarters of a century have elapsed since the first, and still continuing, attempts were made to study that obscure group of phenomena commonly called "psychical", in a "spirit of exact and unimpassioned enquiry". The words were those of the founders of the Society for Psychical Research, and the date was 1882. Few societies in any field of knowledge have obtained such strong support from men and women accustomed to other scientific and scholarly disciplines. It is sufficient to name only a few—Richet, Crookes, Lodge, J. J. Thomson, W. Bateson, Freud, Jung, Barrett, W. James, and Professor and Mrs. Henry Sidgwick. The list could be extended indefinitely.

The evidence studied by the founders of the Society and by their successors can be conveniently considered under four categories. Extensive accounts of the relevant material are readily obtainable in the literature. The four categories are:

1. *Spontaneous cases*, such as those collated and closely analysed by Gurney, Myers and Podmore (1886), and by Mrs. Sidgwick (1922). An example of a spontaneous psychical experience, fairly well authenticated as such cases go, is that of Mrs. Jeanie Gwynne-Bettany, who reported that in her childhood, "I was walking in a country lane at A., the place where my parents then resided. I was reading geometry as I walked along . . . when, in a moment, I saw a bedroom known as the White Room, and upon the floor lay my mother, to all appearances dead. . . . I could not doubt that what I had seen was real, so, instead of going home, I went at once to the house

of our medical man. . . . He at once set out with me for my home, on the way putting questions I could not answer, as my mother was to all appearances well when I left home. I led the doctor straight to the White Room, where we found my mother actually lying as in my vision. . . . She had been seized suddenly by an attack at the heart."

2. *Mediumistic material*, such as that obtained in the investigations of Piper (Lodge, 1890), Leonard (Radclyffe-Hall and Troubridge, 1919) and Cooper (Soal, 1925). The mention of mediumistic utterances at this symposium should be taken to imply *only* paranormal cognition (i.e. telepathy and clairvoyance) involving only living persons. The best of the evidence is impressive and not easy to dismiss.

3. *Qualitative experiments*. These differ from the two preceding categories in the fact that the *material* for study or transmission is prepared by the researcher in advance of the experiment. The weakness of such investigations is that the material, generally drawings, solid objects, verses of poetry and the like, is inaccessible to statistical evaluation. Some important accounts of such work were prepared by Guthrie and Birchall (1883), Guthrie (1885) and Miles and Ramsden (1907 and 1914).

4. *Quantitative experiments*. These appear to have been initiated by the physiologist Charles Richet (1884). Quantitative research forms the main subject matter of this symposium. The leading experiments have been summarized and discussed by J. G. Pratt and others (1940) and by S. G. Soal and F. Bateman (1954).

Within these four categories the printed evidence is ample and sufficient, I believe, to sustain a strong case for the reality of paranormal cognition ("extrasensory perception"). This is common ground to most psychical researchers, though they differ extremely in their judgment as to the quality of individual cases. It is also commonly held by psychical researchers that the facts deserve greater recognition from practitioners in other fields of scientific work than they have thus far been accorded.

Scientific recognition is a term not easy to define. Scientific societies do not adopt resolutions of approval of new discoveries, for this might be the road to scientific dogmatism. Rather, new research findings grow into acceptance gradually, in most instances. This must be especially so if the claims are of a highly novel character. Such widely different things as the anaesthetic use of chloroform and G. Cantor's theory of infinite numbers were accepted only after prolonged and bitter struggle. On the question of psychical research, after three-quarters of a century the disapproval, indifference or apathy of scientific men seem "unconscionably long a-dying".

Psychology is the field in which one would expect to find the greatest amount of interest and study of this subject. In 1938 and again in 1952 L. Warner (Warner and Clark, 1938; Warner, 1952) sought by means of a questionnaire which he issued to several hundred Fellows of the American Psychological Association to obtain information about that question and some others. He found, in both cases, that of those who offered their opinions on extrasensory perception, only one-fifth arrived at their views after reading the published experimental reports. He found also that over the period of fourteen years there was very little change favourable to belief in extrasensory perception. In both surveys, three-quarters of the respondents believed no more than that extrasensory perception was either "a remote possibility" or "merely an unknown".

Having given some indication of the strength and high quality of the evidence in psychical research, I now turn to the reasons why members of other scientific professions appear to react to the subject with feelings of doubt, indifference or aversion. I shall divide the reasons into four classes:

1. The seeming irrelevance of the subject to other sciences.
2. The lack of any repeatable experiment.
3. Disagreement among psychical researchers as to the quality of much of the evidence.
4. Claims made that are unsupportable by the published evidence.

1. The seeming irrelevance of the subject to other sciences

Few scientific men can find any meaningful place for paranormal cognition in their own fields. Sir Oliver Lodge supposed there might be a link between the physical and psychical worlds by means of what was then known as the ether of space. His conjectures have obtained little support. Dr. J. B. Rhine has spoken of extrasensory perception and psychokinesis as giving new meaning to the study of biology, medicine, psychology, "the government of man", "the freedom of the individual" and the struggle against physicalistic Communism. He has also projected a new "Copernican revolution" as a logical consequence of the experiments in psychokinesis (dice-throwing). Those views are perhaps a little speculative. More instructive and interesting are the cautious conjectures of A. C. Hardy (1953) in respect of possible relations between the findings of psychical research and evolutionary theory. Some psychiatrists have written of the significance of paranormal occurrences to their own studies. Something may come of this, but the evidence from this source is slight and often ambiguous. It has been said—but the reference has evaded recent research—that the first mention of Freud in any English writing appeared in a paper by a psychical researcher, Frederic Myers, in the Proceedings of the Society for Psychical Research.

The name of Myers recalls to us that the main title of his classic of psychical investigation was *Human Personality* (Myers, 1903). Psychical research is the study of personality, though this fact is sometimes lost sight of. If we try to look forward, it may seem very probable that the interrelations of paranormal cognition with other factors of personality such as extraversion-introversion, dominance-submission, intelligence, neurotic tendencies and the like, will form the main line of advance in psychical research in the coming years. We may learn something.

Yet by and large it is difficult not to feel some sympathy with those engaged in experimental and applied psychology

who have thus far detected little of interest in our work that has any bearing on theirs. No doubt time will change the scene, but progress is slow. In brief the subject at the present stage of its difficult history seems to have no immediate relevance to the subject matter of other sciences.

2. Failure to produce a repeatable experiment

The major difficulty of qualitative research—that of estimating the chance factor—was overcome in quantitative research by application of the calculus of probability. This was an important step forward.

The second advantage, characteristic of all scientific work in which quantitative methods are used, is the opportunity they give to create repeatable experimentation. By this is meant the designing of an experimental set-up which, found in practice to produce a significant effect, can be repeated by any competent person at any time in the foreseeable future with approximately similar significant results. After thirty years, psychical researchers have failed to produce one repeatable experiment. Yet more than sixteen years have passed since R. A. Fisher (1938) made the following comment on extrasensory perception experiments: "Perhaps I may say, with respect to the use of statements of very long odds, that I have before now, criticised their cogency on the grounds, not only that the procedure of calculation is often questionable, but that they are much less relevant to the establishment of the facts of nature than would be a demonstration of *the reliable reproducibility of the phenomena*".

The failure of psychical research to meet the fundamental inductive principle of science was bound to lead to embarrassing questions. Mr. Spencer Brown has raised one such question. There may be others. I shall return to the pressing need for repeatability at the close of this paper.

3. Disagreement on the quality of the printed evidence

It might be supposed by the friendly inquiring scientist that even if psychical researchers are unable to obtain repetition of their experimental phenomena, at any rate they presumably agree on the validity of those experiments that have not been susceptible to repetition. By this I mean that there are perhaps certain experiments reported in the literature, call them A, B, C, D, E, etc., on which psychical researchers are of the unanimous opinion that *those* experiments certainly cannot be explained away on "normal" hypotheses. It is the purpose of this section of the paper to point to the discouraging fact that psychical researchers of undoubted authority do not agree among themselves as to whether some of the leading experiments are conclusive evidence for paranormality.

To show how real this disagreement is it is necessary only to consider part of the contents of the two most important surveys of the evidence published in recent times. The first is *Extrasensory Perception After Sixty Years* (J. G. Pratt and others, 1940), written by five members of the Parapsychology Laboratory at Duke University. In further references I shall abridge the title of this book to *ESP Sixty* The other important work is *Modern Experiments in Telepathy* (S. G. Soal and F. Bateman, 1954). This title will be abridged hereafter to *Modern Experiments.*

In *ESP Sixty* the authors discussed thirty-five "counter-hypotheses" which they rightly held must be met before the results of an ostensibly psychical experiment could be accepted as valid evidence for paranormal cognition. Those counter-hypotheses included such contingencies as non-random shuffling of cards, sensory leakage in the experimental room, incompetence of the experimenters, errors of recording the targets and calls, dubious statistical methods and so on.

The authors examined 142 published reports, and of these only six survived the counter-hypotheses. The claims made for those six experiments will now be compared with the opinions on the same experiments presented in *Modern Experiments*, a book which it is necessary to note is described by the authors as "a fairly detailed account of the better class of experiments in extrasensory perception which have been carried out in the past thirty-five years". The six experiments of *ESP Sixty* all happen to be covered by the time period mentioned. We may now turn to them.

(i) *Pratt-Woodruff*, the title (being the experimenters' names) by which it is widely known in the field, was a test of sixty-six subjects by a procedure known as "screened-touch-matching". There were about 60,000 trials, the total score was measurably above chance, χ, the normal deviate, being 4·99, representing odds against the null hypothesis of the order of a million to one. I should regard this research as an example of experimental rigour applied to the last degree of scientific necessity. But the authors of *Modern Experiments* judge it to be only "a fairly good series", an unexpected comment that has "surprised" Dr. McConnell (1954). Hence, we have here four opinions, three unreservedly favourable to the research, one (from a most distinguished source) not so favourable.

(ii) *Lucien Warner's "A Test Case"*. This was a series of 250 trials. The subject, a woman, was accommodated in a ground-floor room. The two experimenters (described as "psychologically trained") were in a locked room on the first floor and in the opposite wing of the house to that in which the subject was placed. The cards were drawn singly from a pack which was shuffled after each trial. There was no conversation between the experimenters, and the signalling system was one-way; from subject to experimenters only. The average score was 9·3 per run, a remarkable result if genuine; mean chance expectation being the usual 5·0 per run. The normal deviate was 6·8, and the odds against the null hypothesis too astronomically remote to be worth computing. This truly astonishing result obtained under seemingly impeccable conditions is treated by Soal and Bateman with total silence, for it is nowhere mentioned in their most thorough survey of thirty-five years' research.

(iii) *Pearce-Pratt.* This well-known clairvoyance investigation was carried out with the principal experimenter, J. G. Pratt, and the subject, Hubert Pearce, accommodated in separate buildings at two distances apart, 340 feet and 740 feet. The number of trials was 1850, the average score per run was 7·53, and the normal deviate was 10·87, which has an associated probability of 10^{-20}.

The Pearce-Pratt series is warmly commended in *Modern Experiments.*

(iv) *Turner-Ownbey.* This was a short series of telepathy experiments in which the agent, Miss Ownbey, and the subject, Miss Turner, were at first 250 miles and later 300 miles apart. Each day's work comprised twenty-five trials, and (what is important to note) it was arranged that they would despatch their records to J. B. Rhine at the end of each day's work. For the first three days, however, Miss Turner failed to carry out this instruction, sending her records instead to Miss Ownbey. The results claimed under those regrettable conditions were phenomenal, the scores for the three runs being 19, 16 and 16. In the subsequent experiments, in which Miss Turner faithfully followed the prescribed rule, her scores were in close accord with mean chance expectation.

Modern Experiments, if I understand the authors correctly, is on the whole, favourable to the Turner-Ownbey affair, though some doubt is expressed. The authors say, "we shall have to accept the results of the experiment, or assume that the two experimenters were in collusion".

(v) *Riess.* The experimenter, Dr. B. F. Riess, was a psychologist. The subject (name not disclosed) was a young woman music teacher. Whilst Riess exposed the cards in his home, the subject wrote down her guesses at her own residence more than quarter of a mile away. There were seventy-four runs (of the customary twenty-five trials per run). The average score per run was the extraordinary figure of 18·24, which is the largest average score in the history of the subject. In one spell of nine runs all the scores were more than 20. The normal deviate works out at 53·57, which is the largest on record for a single series. The probability against the null hypothesis appears to be about 10^{-700}.

The case is accepted by the Duke group as one of the six most conclusive series in history, but in *Modern Experiments* it is dismissed as one of "the questionable experiments . . . which in the present book we have scarcely considered to be worthy of mention. . . ."

(vi) *Murphy and Taves.* This was a long and careful piece of research involving many subjects. All the necessary precautions were taken. For each subject there were several kinds of tests, the experimental material consisting of playing-cards and various forms of ESP cards. The number of trials was 175,000. The total scores were not significantly above chance expectation, but significant correlations were found between certain of the ESP tasks. These correlations had not been foreseen in the experimental design, and hence their statistical quality is only that of *ad hoc* hypotheses.

As indicated, in *ESP Sixty* the experiment and its outcome are judged to be one of the six most impressive cases on record. In *Modern Experiments* the case is not mentioned.

Summary. Six out of 142 researches were judged by the first group of authors to surmount all plausible counter-hypotheses. Of these cases, the second group of (two) authors assent unhesitatingly to their predecessors' views on one research (or two if the Turner-Ownbey matter be included), are not completely satisfied with a third research, reject a fourth, and ignore the existence of the two remaining researches.

Clearly, there is no unity of opinion among leading psychical researchers as to what constitutes valid evidence.

4. Unsubstantiated Claims

This is the fourth, and last, of the factors which appear to be operating against the scientific recognition of parapsychology. The question now refers to certain exaggerated and unsubstantiated claims that appear with uncomfortable frequency in the literature of parapsychology. To deal with the problem comprehensively would call for the writing of several papers. Here I shall try only to illustrate the gravity of the matter by taking examples from several pages of one book.

The subject of discourse is psychokinesis, a word that is commonly abridged to "PK". Psychokinesis has been formally defined as: "The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation", or, very colloquially, "the power of mind over matter". The physical objects commonly used in the many reported experiments are ordinary commercial dice. These are caused to be thrown whilst the subject "wills" which face will turn up, and the results are assessed statistically. Dice throwing raises a host of awkward problems, which will not be discussed now; but it is necessary to recognise that all dice are in varying degrees biased, or must be assumed so.

That extrasensory perception is a well-established fact of nature, few who have studied the literature with care and without prejudice will doubt. The evidence for psychokinesis is less satisfactory. It is interesting to note, however, that according to one important authority (Rhine, 1947), "PK implies ESP, and ESP implies PK. . . . Each phenomenon follows from and leads into the other; each derives confirmatory support from the work which established the other." This passage is from *The Reach of the Mind* by J. B. Rhine. The book is not easy to classify. It was regarded as sufficiently suitable for the general public to be published in an abridged form in *Readers Digest.* On the other hand, its pretensions to being classified as a scientific statement may be gleaned from the fact that the author recommended it to the attention of the New York Academy of Sciences (Rhine, 1950). On meeting scientists engaged in other fields of work one commonly discovers that they rarely have time to read the original reports in our subject. So it is convenient for them to have available in a couple of hundred pages a summary of the facts —or the supposed facts. In this book a large number of the pronouncements of claims with respect to psychokinesis evidence are misleading, sometimes they are contrary to the evidence presented in the original reports, and in some places formidable claims are made on the basis of evidence that has never been reported in any scientific paper.

To take one case, it is remarked in the book that two students, H. Hilton and G. Baer, obtained "significantly high scoring in high dice tests". This claim sags somewhat when it is found in the original report that a control series of throws was made to determine the bias of the dice, and that the "high dice" totals in this case were of precisely the same order as in the putatively PK throws. The probability of observed difference between control and supposed PK is a modest 0·27.

With regard to PK work comparing the effect of throwing different numbers of dice at a time, the author of the book endeavours to present evidence (here and elsewhere) that PK is more effective on many dice thrown together than on one or two. For example, in the book it is observed that "the tests with six dice scored higher than those with two dice. The rate with two dice was not much above 'chance', but the results with six were highly significant. This finding is obviously not what one would expect from a physical viewpoint."

These statements are at variance with the experimental facts. In the PK literature two experimental papers cite comparisons of this type. One is the work of Frank Smith (Rhine, 1944), the other is part of the first E. P. Gibson research (Gibson and Rhine, 1943). The results were:

No. of dice thrown	Smith		Gibson	
	Average per run	χ	*Average per run*	χ
2	4·33	3·56	4·80	2·40
6	4·31	2·56	4·07	0·82

(A "run" is twenty-four throws of one die, twelve throws of two, or four throws of six. The expected average, other things being equal, is therefore always 4·00.)

The results for both Smith and Gibson are plainly the exact opposite of those so confidently announced in the book. Moreover, since there was no control on the bias, and the "target" was almost consistently chosen as face 6, the normal deviates computed can have no bearing on the PK hypothesis. One searches the literature in vain for any comparison fitting the book's claim that "tests with six scored higher than those with two dice".

The belief that in psychokinesis experiments the dice do not obey physical laws is constantly urged. H. L. Frick investigated his own psychokinetic powers by throwing at the same time twelve dice described as of medium size and twelve described as small, the former being "approximately twice as large cubically" as the small ones. But, according to the book, "the larger dice gave almost twice as many *hits* above chance average" as compared with the small ones, and therefore "the results obtained were just the opposite of what one might expect from a physical theory".

In face of this seeming modification of the Laws of Motion, a scientific reader might be tempted to consult the original report. With the aid of elementary statistics he would easily find that the discovery reported in the book is dangerously misleading. So closely similar are the scores for the two kinds of dice that the probability of their observed difference is 0·40. Yet the tone of the book is such as to leave the reader to infer that the difference is significant and real. On the basis of such data, and others of little better quality, it is seriously stated on a subsequent page that "the indifference of PK to the physical law of gravity is surely something to ponder".

A factor of distance in relation to psychokinesis is reported to have been introduced in experiments at the Parapsychology Laboratory. "When", observed the author of the book, "the subject was assured that he could do better from the more distant position, the results obtained at twenty-five feet were much better than when he stood close to the dice. After this point was demonstrated, the advantage of distance was played down and that of being close to the dice was emphasized. As a result the relation between the two scoring averages was reversed so that proximity was favoured." These supposed findings, which are later described as "more than suggestive", are clearly of great importance to the scientific world, indicating as they do the abrogation or modification of gravitational law and the existence of a force that can operate invisibly at a distance of twenty-five feet from its source. When the PK reports are searched for further details, we come away empty-handed. No such report exists.

On the strength of dice throwing said to have been performed at Duke University, it is recorded elsewhere that the psychokinetic force is more effective on heavy metal dice than on wooden ones; and also, from the same source, that the shape of the dice—sharp corners, rounded corners, or extremely rounded corners—makes no difference to the power of the human psychic force. Neither of these strange claims can be validated in any of the published reports on psychokinesis. It is mainly on the basis of these and similar unverifiable assertions that the author concludes that "the finding that *mass, number, and form* [of dice] *are not determining conditions in PK tests*, takes its place, then, alongside the discovery that time and space were not limiting factors in ESP". The italics are mine.

Only a few of these rash pronouncements have been quoted in the above paragraphs. Many others could be cited. One wonders what the more objective but friendly type of scientist must think when he is confronted with such highly adorned claims. He might, one surmises, rather easily turn away from psychical research, moved by the uncomfortable realization that a subject in which scientific method and the need for careful reporting are so casually pushed out of the way, is not a field of study in which he would care to indulge.

General summary and conclusions

Four factors operating to produce indifference or hostility to psychical research have been adduced. These are the apparent irrelevance of psychical research to other scientific fields, the failure of psychical researchers to obtain reproducible phenomena, discordant views on the validity of evidence, and inflated claims to discoveries.

These factors are not of equal importance, and it seems clear to me that the most pressing need is that of obtaining repetition of results. Given the discovery of one experimental design with repeatability qualities, however simple and modest it might be, one would expect others to follow in its steps. The conflicts of belief amongst psychical researchers (to say nothing of the scientific world in general) would presumably lose significance, and the temptation to make inflated claims ought to decline. Results that could be predicted, at least to a fair degree of precision, are by definition, *science*, and therefore "respectable", and deserving of recognition. This desirable consummation would be nearer in that situation than it has been in the past. A preceding speaker has expressed the urgent need for experimental repeatability. I can only add my support to that view, and express the hope that the whole strength of research should now be devoted to that end.

REFERENCES

FISHER, R. A. (1938). *J. Parapsychol.*, **2**, 267.
GIBSON, E. P., and RHINE, J. B. (1943). *J. Parapsychol.*, **7**, 118.
GURNEY, E., MYERS, F. W. H., and PODMORE, F. (1886). Phantasms of the Living (2 vols.). London: Trübner.
GUTHRIE, M. (1885). *Proc. Soc. psych. Res., Lond.*, **3**, 424.
GUTHRIE, M., and BIRCHALL, J. (1883). *Proc. Soc. psych. Res., Lond.*, **1**, 263.
HARDY, A. C. (1953). *Proc. Soc. psych. Res., Lond.*, **50**, 96.
LODGE, O. J. (1890). *Proc. Soc. psych. Res., Lond.*, **6**, 443.
MCCONNELL, R. A. (1954). *J. Parapsychol.*, **18**, 252.
MILES, C., and RAMSDEN, H. (1907). *Proc. Soc. psych. Res., Lond.*, **21**, 60.
MILES, C., and RAMSDEN, H. (1914). *Proc. Soc. psych. Res., Lond.*, **27**, 279.
MYERS, F. W. H. (1903). Human Personality and Its Survival of Bodily Death (2 vols.). London: Longmans, Green.
PRATT, J. G., Rhine, J. B., Smith, B. M., Stuart, C. E., and Greenwood, J. A. (1940). Extra-sensory Perception After Sixty Years. New York: Henry Holt.
RADCLYFFE-HALL and UNA, LADY TROUBRIDGE (1919). *Proc. Soc. psych. Res., Lond.*, **30**, 339.
RHINE, J. B. (1944). *J. Parapsychol.*, **8**, 287.
RHINE, J. B. (1947). The Reach of the Mind. New York: Wm. Sloane.
RHINE, J. B. (1950). *Trans. N.Y. Acad. Sci.*, Series II, Vol. 12, 164.

RICHET, C. (1884). *Rev. Philos.*, Dec.
SIDGWICK, E. H. (1922). *Proc. Soc. psych. Res., Lond.*, **33**, 1.
SOAL, S. G. (1925). *Proc. Soc. psych. Res., Lond.*, **35**, 471.
SOAL, S. G., and BATEMAN, F. (1954). Modern Experiments in Telepathy. London: Faber & Faber.
WARNER, L. (1952). *J. Parapsychol.*, **16**, 284.
WARNER, L., and CLARK, C. C. (1938). *J. Parapsychol.*, **2**, 296.